Identification of glycosylation sites in the SU component of the Avian Sarcoma/Leukosis virus Envelope Glycoprotein (Subgroup A) by mass spectrometry

Mamuka Kvaratskhelia, a Patrick K. Clark, b Sonja Hess, c Deborah C. Melder, d Mark J. Federspiel, d and Stephen H. Hughes*, e

a Center for Retrovirus Research and Comprehensive Cancer Center, College of Pharmacy, The Ohio State University Health Sciences Center, Columbus, OH 43210, USA
b Basic Research Program, SAIC-Frederick, Frederick, MD 21702, USA
c National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892, USA
d Molecular Medicine Program, Mayo Clinic College of Medicine, Rochester, MN 55905, USA
e HIV Drug Resistance Program, Center for Cancer Research, NCI-Frederick, Frederick, MD 21702, USA

Received 20 April 2004; accepted 18 May 2004

Abstract

We used enzymatic digestion and mass spectrometry to identify the sites of glycosylation on the SU component of the Avian Sarcoma/Leukosis virus (ASLV) Envelope Glycoprotein (Subgroup A). The analysis was done with an SU(A)-rIgG fusion protein that binds the cognate receptor (Tva) specifically. PNGase F removed all the carbohydrate from the SU(A)-rIgG fusion. PNGase F is specific for N-linked carbohydrates; this shows that all the carbohydrate on SU(A) is N-linked. There are 10 modified asparagines in SU(A) (N17, N59, N80, N97, N117, N196, N230, N246, N254, and N330). All conform to the consensus site for N-linked glycosylation NXS/T. There is one potential glycosylation site (N236) that is not modified. Removing most of the carbohydrate from the mature SU(A)-rIgG by PNGase F treatment greatly reduces the ability of the protein to bind Tva, suggesting that carbohydrate may play a direct role in receptor binding.

Keywords: Avian; Retrovirus; Envelope; Glycosylation; Tva

Introduction

Retroviruses infect susceptible cells via an interaction between the virally encoded envelope glycoproteins and cognate receptors on the surface of the cell. This interaction causes structural changes in the envelope glycoprotein that lead to the fusion of the viral membrane and the cellular membrane; membrane fusion places the viral core in the cytoplasm of the cell. The Avian Sarcoma Leukosis Virus (ASLV) envelope glycoproteins are a group of related retroviral envelope glycoproteins that have been named based on the specific receptors they interact with and their reaction with antibodies. The main ASLV envelope subgroups are A–E. The differences in the envelopes that allow the different ASLV envelopes to interact with their cognate receptors map to defined segments of the envelope (hr1, hr2, vr3) (Bova et al., 1986; Dorner et al., 1985). The subgroup (A) receptor is Tva, a small protein of unknown function that is related (by sequence and structure) to the ligand binding repeat of the low-density lipoprotein receptor (LDLR) (Bates et al., 1993). The B, D, and E viruses interact with different aspects of an integral membrane protein, originally called CAR1, now called Tvb, that is a member of the tumor necrosis factor receptor (TNFR) family (Adkins et al., 1997; Brojatsch et al., 1996). The subgroup C receptor has not been identified. Appropriately folded Tva can be expressed in soluble form; this makes it easy to study the interactions of the envelope glycoprotein and the receptor. Although Tvb is an integral membrane protein, a soluble peptide (15 aa long) can substitute for...
TvB with respect to its interactions with subgroup B envelope.

Retroviral envelopes (including the ASLV envelopes) are translated from a spliced message. The proteins are glycosylated and cleaved by host proteases to produce mature proteins with two components: surface (SU) and transmembrane (TM). The mature protein is trimeric; TM anchors the envelope glycoprotein in the viral membrane. TM is the portion of the envelope that is directly involved in the fusion of the viral membrane and the membrane of the target cell during infection. Although the details of this fusion event are not completely understood, major rearrangements of TM embed the N-terminal fusion peptide of TM in the cellular membrane and subsequent rearrangements of TM bring the cell membrane and the viral membrane together, leading to fusion.

The initial rearrangement of TM, in which the fusion peptide is embedded in the membrane of the host cell, depends on changes in the structure of SU (see Coffin et al., 1997, for review). In the simplest mechanism, exemplified by the Moloney Murine Leukemia Virus (MLV), when SU binds to its cognate receptor on the surface of the cell, this interaction triggers a change in the structure of SU, which in turn leads to changes in the structure of TM. Human Immunodeficiency Virus Type 1 (HIV-1) envelope interacts with both a receptor (CD4) and a co-receptor (commonly CXCR4 or CCR5). HIV-1 envelope interacts first with CD4, which alters the structure of the envelope and allows it to bind to the co-receptor. It is the interaction with the co-receptor that triggers the critical structural change in SU leading to the rearrangement of TM and fusion of the cellular and viral membranes. However, for both MLV and HIV-1, pH does not appear to play any direct role in the events that lead to membrane fusion. In this sense, the reactions carried out by retroviral envelopes, which are mechanistically similar to the reactions carried out by the HA protein of the flu virus, are differently controlled. In the flu system, the binding of the SU-equivalent component (HA1) to the receptor serves to bring the virus into an acidic compartment in the cell where the pH change triggers the change in the structure of the TM-equivalent portion (HA2); in the MLV and HIV envelopes, it is the binding of SU that triggers the structural change and the virus is believed to fuse to the outer membrane of the target cell without being transported into an acidic compartment within the cell.

The mechanism that controls the fusion of the envelope glycoprotein of the Avian Sarcoma/Leukosis viruses (ASLVs) is controversial. There is good agreement that binding the receptor to SU is an essential step, and that there is some sort of structural change in SU when the receptor binds. The question is whether binding is both necessary and sufficient to trigger all the steps that lead to membrane fusion. Mothes et al. (2000) presented evidence that additional structural changes are required and that these changes depend on low pH. In this model, the mechanism that controls ASLV envelope fusion is in some ways intermediate between the classical retroviral mechanism (which depends only on binding the receptor to induce the structural change) and the flu mechanism (which depends only on low pH to induce the structural change). However, the Mothes et al. model has been challenged; Earp et al. (2003) have argued that low pH does not play a critical role in ASLV envelope fusion.

For ASLV envelope, it is generally agreed that binding the receptor at neutral pH causes a significant change in the conformation of the protein (Damicco et al., 1998; Earp et al., 2003; Gilbert et al., 1995; Hernandez et al., 1997). This change exposes the TM fusion peptide, which then associates with the membrane of the target cell. The next step is the fusion of the viral and cellular membranes, which is associated with TM rearranging to form a stable six-helix bundle. The controversy is whether this step does, or does not, depend on low pH. Earp et al. (2003) have suggested that this process proceeds, at least to the stage of hemifusion of the membranes, at neutral pH. However, there are good reasons to believe low pH plays a critical role in this process.

Earp et al. (2003) suggested that some later step in the infection process, perhaps uncoating, is pH dependent. However, Smith et al. (2004) showed that receptor binding and low pH are both required to cause the efficient formation of stable TM trimers; receptor binding alone produces relatively small amounts of stable TM trimers. Melikyan et al. (2004) presented evidence that low pH is required to reach hemifusion of the membranes. Moreover, if HIV-1 particles carrying a Vpr-β lactamase fusion are pseudotyped with ASLV subgroup A envelope, delivery of the viral core into the cytoplasm is blocked by treating the cells with NH4Cl. The effect of NH4Cl on the entry of HIV-1 particles is also seen when the particles are pseudotyped with VSV-G, which is pH dependent. However, if HIV-1 envelope (which is not pH dependent) is used, entry is insensitive to NH4Cl (J. Young, personal communication). These results make the ASLV envelope glycoproteins particularly interesting targets for exploring the relationship between protein structure and membrane fusion.

Although there are partial structures for the SU components of HIV-1 envelope and MLV envelope, there is no corresponding information for the SU proteins of the ASLV viruses. ASLV SU(A) is, like other retroviral SU proteins, heavily glycosylated. As part of a larger effort to understand the structure and function of SU(A), we physically mapped the glycosylation sites in SU(A) by a combination of enzymatic digestion and mass spectrometry. This approach has been used to characterize the glycosylation patterns of many glycoproteins, including HIV-1 envelope (Borchers and Tomer, 1999; Dell and Morris, 2001; Leonard et al., 1990; Zhu et al., 2000). To simplify the isolation and purification of the SU(A) protein, we used an SU(A) fused to rabbit immunoglobulin [SU(A)-rIgG] (see Materials and methods). This fusion
protein binds the receptor and has the appropriate properties. The carbohydrates can be completely removed from denatured SU(A) by PNGase F demonstrating that only N-linked glycosylation is present. Of the 11 asparagines in SU(A) that are, by sequence, sites where carbohydrates could be attached, 10 were glycosylated. PNGase F treatment of intact, unmodified SU(A)-rIgG greatly reduces its ability to interact with Tva, suggesting that carbohydrate

Fig. 1. Immunoblot and pull-down analysis of the SU(A)-rlgG and chicken sTva-mlgG proteins. The SU(A)-rlgG and sTva-mlgG proteins are indicated on the right. The volumes (in microliters) of cell culture supernatants used and any glycosidase treatment the sample received are given at the top of each immunoblot. Molecular sizes (in kilodaltons) are given on the left. (A) First, either the SU(A)-rlgG protein was immunoprecipitated from cell culture supernatant with anti-rabbit agarose beads or the sTva-mlgG was immunoprecipitated with anti-mouse agarose beads. The bound complexes were washed and then were mixed with cell culture supernatant containing either the SU(A)-rlgG or sTva-mlgG protein or mock supernatants. After an additional wash, the proteins were denatured, separated by SDS-PAGE, and transferred to a nitrocellulose membrane. The Western immunoblot was probed with peroxidase conjugated anti-rabbit and anti-mouse IgG, and the bound protein–antibody complexes were visualized by chemiluminescence. (B) The SU(A)-rlgG protein was immunoprecipitated from cell culture supernatant with anti-rabbit agarose beads, and the bound complexes were washed. The samples were treated with Endo H or PNGase F and then washed. The bound proteins were mixed with cell culture supernatant containing sTva-mlgG protein or mock supernatant, and the bound complexes were subsequently washed. The Western immunoblot was prepared and probed as in (A).
may play a direct role in the interactions of the SU(A) envelope and Tva.

**Results**

*Phenotype of the SU(A)-rIgG protein*

The SU(A)-rIgG fusion protein is soluble and the rabbit IgG portion allows a simple purification of biologically active ASLV SU(A) in good yield from chicken cells that express the fusion protein (see Materials and methods). We have previously demonstrated that susceptible chicken cells expressing the SU(A)-rIgG protein are specifically resistant to ASLV(A) infection by receptor interference but remained susceptible to infection by other ASLV subgroups (Holmen and Federspiel, 2000). The SU(A)-rIgG protein migrates as a relatively compact band of approximately 82 kDa on SDS-PAGE as determined by Western immunoblot (Fig. 1A). For the same reasons that it is convenient to use the SU(A)-rIgG fusion protein, it is convenient to use a corresponding fusion protein in which Tva is linked to mouse immunoglobulin (sTva-mIgG). This sTva-mIgG fusion has been used to estimate that the binding affinity of chicken Tva for ASLV(A) envelope glycoprotein expressed on chicken DF-1 cells is approximately 0.5 nM (Melder et al., 2003). Chicken sTva-mIgG binds the SU(A)-rIgG protein efficiently in a pull-down assay (Fig. 1A). We treated the SU(A)-rIgG to test whether removing the carbohydrate present on the SU(A)-rIgG protein affected its ability to bind to sTva-mIgG. Native SU(A)-rIgG protein was bound to agarose beads through the IgG component and treated with either Endo H, a glycosidase that cleaves high mannose structures from N-linked glycoproteins, or PNGase F, a glycosidase that cleaves N-linked glycan chains from glycoproteins. PNGase F cleaved a significant portion of carbohydrate from the SU(A)-rIgG protein; after PNGase F treatment, SU(A)-rIgG migrated at position indicating a mass of approximately 65 kDa, while Endo H removed only a small amount of carbohydrate (Fig. 1B). The expected MW of the protein component of SU(A)-rIgG is 62764 Da; this suggests PNGase F removes the majority of the carbohydrate from the properly folded native protein. The ability of the deglycosylated SU(A)-rIgG to bind sTva-mIgG was determined using a pull-down assay. The Endo H-treated SU(A)-rIgG was efficiently bound to the sTva-mIgG receptor; the PNGase F-treated SU(A)-rIgG did not bind (Fig. 1B). We conclude from these experiments that the SU(A)-rIgG protein is biologically active and that at least some of the N-linked carbohydrate is required for efficient interaction with the Tva receptor.

*SU(A)-rIgG analysis by mass spectroscopy*

The experimental strategy for identifying the glycosylation sites in SU(A)-rIgG is depicted in Fig. 2. First, the protein was denatured and the cysteine residues were modified.
SU(A) contains 14 cysteine residues, the majority of which are thought to be involved in disulfide bonds that help define the architecture of the protein. SU(A)-rIgG was unfolded and the cysteines modified with acrylamide to ensure efficient deglycosylation and proteolysis of the protein. In addition, the peptides that contained acrylamide modified cysteines were readily identified by mass spectrometry (MS). Matrix assisted laser desorption/ionization time-of-flight (MALDI-ToF) data for the full-length deglycosylated protein and MS and MS/MS data for the proteolytic fragments indicated that all cysteine residues were fully modified under our reaction conditions. The total mass of the glycan modifications could be estimated by comparing full-length glycosylated and deglycosylated SU(A)-rIgG (Fig. 2). The individual glycosylation sites in the protein were determined by comparative MS and MS/MS analyses of proteolytic fragments derived from the glycosylated and deglycosylated proteins.

MALDI-ToF was used to determine the MW of glycosylated SU(A)-rIgG. Fig. 3A shows that the glycosylation is somewhat heterogeneous. Two main peaks with MW of 86510 and 87612 Da were detected. Subtracting the calculated MW of the SU(A)-rIgG polypeptide chain containing acrylamide-modified Cys residues (adding acrylamide residues to the cysteines increased the MW to 64184 Da) from these experimental values showed that there is approximately 23 kDa of carbohydrate. The denatured SU(A)-rIgG was treated with PNGase F, which specifically removes N-linked carbohydrate chains. Fig. 3B shows that PNGase F-treated SU(A)-rIgG produced a single peak of 64194 Da. This value is within ±10 Da (0.01% mass accuracy) of the calculated MW of the SU(A)-rIgG polypeptide chain. Thus, PNGase F treatment removed all glycans indicating that all of the glycosylation sites in SU(A)-rIgG involve asparagine residues.

The locations of the glycosylated asparagines were determined by subjecting glycosylated and deglycosylated proteins to proteolysis in parallel. This generated small peptide fragments that were used for MS/MS analyses. The proteins were digested using trypsin, chymotrypsin, or GluC. The peptide mixtures were analyzed with MALDI-ToF and electrospray ionization time-of-flight (ESI-ToF) mass spectrometry. Representative MALDI-ToF data are shown in Fig. 4. Comparison of the peptides from glycosylated and deglycosylated SU(A)-rIgG showed that some peptide fragments were present in digests from both protein preparations. These peptides did not contain sugar residues. However, several proteolytic fragments were detected only in the PNGase F-treated samples (Fig. 4A). These peptides were generated as a result of the enzymatic removal of the N-linked oligosaccharide and were subsequently assigned to specific SU(A) sequences.

The deglycosylated peaks had a distinctive feature. PNGase F treatment converts the glycosylated Asn to Asp (plus 1 Da mass difference). We were able to detect this difference in MW in both MS and MS/MS analyses. Based on nucleic acid sequence, SU(A)-derived peptides 16–24, 53–57, 55–63, 58–79, 64–84, 68–87, 95–110, 111–134, 219–234, and 229–241 (see Table 1) should contain a single Asn residue; each of these fragments exhibited a mass increment of about 1 Da associated with the Asn to Asp conversion. In contrast, peptides 1–18, 191–202, 193–204, 219–237, and 304–332 (see Table 1) were predicted to contain two Asn residues showed a mass difference of plus 1 Da, suggesting that only one of the Asn residues was glycosylated. MS/MS analyses of these peptide fragments with ESI-ToF were used to locate the glycosylation sites. Representative MS/MS data on this peptide confirmed that both Asn246 and Asn254 were converted to Asp upon deglycosylation of the protein with PNGase F. One chymotryptic peptide (242–270) that should contain two Asn residues displayed a mass increment of approximately 2 Da in MS analyses (Table 1). The MS/MS data on this peptide confirmed that both Asn246 and Asn254 were converted to Asp upon PNGase F treatment. Thus, we have been able to accurately identify all the glycosylated asparagines residues in SU(A) (Table 1).

Fig. 3. MALDI-ToF of full-length glycosylated (A, upper panel) and deglycosylated (B, lower panel) SUA-IgG. Sinapinic acid was used as a matrix. The glycosylated protein exhibited two major peaks at 86510 and 87612 Da indicating heterogeneous glycosylation. PNGase F treatment of SU(A)-IgG produced a single peak with a MW of 64194 Da. This value is within ±10 Da (0.01% mass accuracy) of the calculated MW of the full-length polypeptide chain of SU(A)-IgG, which means that PNGase F removed all the sugar residues.

Table 1. MS/MS data of chymotryptic peptide 193–204 are depicted in Fig. 5. These data indicate that Asn196 and not Asn193 was converted to Asp upon deglycosylation of the protein with PNGase F. One chymotryptic peptide (242–270) that should contain two Asn residues displayed a mass increment of approximately 2 Da in MS analyses (Table 1). The MS/MS data on this peptide confirmed that both Asn246 and Asn254 were converted to Asp upon PNGase F treatment. Thus, we have been able to accurately identify all the glycosylated asparagines residues in SU(A) (Table 1).
A summary of the MS and MS/MS analyses in the context of SU(A) primary structure is given in Fig. 6. The use of three proteolytic enzymes (trypsin, chymotrypsin, and GluC) was complementary and enabled us to obtain greater than 92% coverage of the amino acid sequence of SU(A). MALDI-ToF analyses of the PNGase F-treated protein indicated that SU(A) contains only N-linked glycans (Fig. 3). Importantly, all 16 asparagines of SU(A) were amenable to MS and MS/MS analyses enabling us to accurately identify glycosylated and nonglycosylated residues (Fig. 6). In SU(A), the following 10 sites are glycosylated: N17, N59, N80, N97, N117, N196, N230, N246, N254, and N330. All of the asparagines that carry carbohydrate conform to the consensus sequence NXS/T. Of the 11 NXS/T sites in SU(A), 1 (N236) is not glycosylated.

**Discussion**

Mass spectrometry shows that SU(A)-rIgG contains approximately 23 kDa of carbohydrate. All the carbohydrate is N-linked. There are 11 potential N-linked glycosylation sites in SU(A); 10 of the 11 sites have a carbohydrate modification. The one potential N-linked glycosylation site that is not modified in SU(A) is N236. The S at 238 can be mutated to A (which would have blocked glycosylation of N236 had it been a site for glycosylation) without causing serious disruption in the structure or function of SU(A). Virus carrying the S238A mutation in SU(A) retained 37% of wild-type infectivity (Delos et al., 2002).

Mutating the S or T residues just downstream of three of the N-linked glycosylation sites in SU(A) (N59, N196, and N254) affected the folding and processing of SU(A) (Delos et al., 2002). Multiple mutations at other S or T residues associated with other glycosylation sites also had deleterious effects on SU(A). However, the mutagenesis experiments and our PNGase F digestion experiments ask somewhat different questions. The mutagenesis experiments showed which sites (more specifically the S or T residues in these sites) are essential. These sites may, or may not, be sites where there is an essential carbohydrate. The effects of the SU(A) mutations may not be a direct effect of blocking glycosylation; the primary effect of the mutations might involve the folding and processing of the protein. Our PNGase F experiments involve the removal of carbohydrate from a mature, properly folded.
protein and suggest that carbohydrate might play a direct role in binding SU(A) to Tva.

What is surprising is not that carbohydrate could have a role in binding Tva, but that all asparagines that carry the carbohydrate in SU(A) are conserved in the various ASLV envelope subgroups. This suggests that the various ASLV SU proteins will have a common pattern of carbohydrate modification. This makes good sense if the carbohydrate plays a structural role; however, it does raise a question: Why would a carbohydrate group that is present in all the

Table 1  
Summary of deglycosylated peptides identified with ESI-ToF MS and MS/MS

<table>
<thead>
<tr>
<th>Exp. peptide mass (charge state)</th>
<th>M_r exp. (Da)</th>
<th>M_r calc. (Da)</th>
<th>Delta (Da)</th>
<th>Amino acid sequence*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chymotryptic peptides</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>505.72  (2+)</td>
<td>1009.44</td>
<td>1008.46</td>
<td>0.98</td>
<td>16-A(N/D)RTGQTDF-24</td>
</tr>
<tr>
<td>1025.45  (1+)</td>
<td>1024.45</td>
<td>1023.46</td>
<td>0.99</td>
<td>55-VSDT(N/D)C*TTL-63</td>
</tr>
<tr>
<td>1056.47  (2+)</td>
<td>2110.94</td>
<td>2109.98</td>
<td>0.94</td>
<td>64-GTDLVSSADFTGGPD(N/D)STTL-84</td>
</tr>
<tr>
<td>956.97  (2+)</td>
<td>1911.94</td>
<td>1910.98</td>
<td>0.97</td>
<td>95-KL(N/D)VSMWDEPEPQTLI-110</td>
</tr>
<tr>
<td>1203.59  (2+)</td>
<td>2405.19</td>
<td>2404.21</td>
<td>0.98</td>
<td>111-GSQSLP(N/D)TNTIAQSGTGGG<em>C</em>VGF-134</td>
</tr>
<tr>
<td>767.85  (2+)</td>
<td>1533.70</td>
<td>1532.66</td>
<td>0.94</td>
<td>193-NMY(N/D)C*SQVQRQY-204</td>
</tr>
<tr>
<td>761.32  (2+)</td>
<td>1520.64</td>
<td>1519.68</td>
<td>0.94</td>
<td>229-V(N/D)QSQEINESEF-241</td>
</tr>
<tr>
<td>1015.82  (3+)</td>
<td>3044.46</td>
<td>3042.45</td>
<td>2.01</td>
<td>242-SFTV(N/D)C<em>TASSLG(N/D)ASGC</em>C*GKAGTILPGKW-270</td>
</tr>
<tr>
<td>Tryptic peptides</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1082.03  (2+)</td>
<td>2162.06</td>
<td>2161.11</td>
<td>0.95</td>
<td>1-DVHLLEQPGNLWITWA(N/D)R-18</td>
</tr>
<tr>
<td>837.84  (2+)</td>
<td>1673.68</td>
<td>1672.74</td>
<td>0.94</td>
<td>53-GYVSDT(N/D)C*TTLGTDPR-67</td>
</tr>
<tr>
<td>1051.98  (2+)</td>
<td>2101.96</td>
<td>2101.00</td>
<td>0.96</td>
<td>68-LVSSADFTGGPD(N/D)STTLTYR-87</td>
</tr>
<tr>
<td>788.82  (2+)</td>
<td>1575.65</td>
<td>1574.68</td>
<td>0.97</td>
<td>191-FWNNMY(N/D)C*SQVQR-202</td>
</tr>
<tr>
<td>388.21  (2+)</td>
<td>774.42</td>
<td>773.44</td>
<td>0.98</td>
<td>327-VLV(N/D)SSR-333</td>
</tr>
<tr>
<td>GluC peptides</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>767.37  (3+)</td>
<td>2299.11</td>
<td>2298.05</td>
<td>0.98</td>
<td>58-T(N/D)C*TTLGTDRLVSSADFTGGPD-79</td>
</tr>
<tr>
<td>654.65  (3+)</td>
<td>1960.95</td>
<td>1959.98</td>
<td>0.97</td>
<td>219-IQC<em>C</em>TRRGGKWV(N/D)QSQE-234</td>
</tr>
<tr>
<td>773.38  (3+)</td>
<td>2317.14</td>
<td>2316.15</td>
<td>0.99</td>
<td>219-IQC<em>C</em>TRRGGKWV(N/D)QSQEINE-237</td>
</tr>
</tbody>
</table>

* N/D conversion identified by MS/MS.
* Cysteine modified with acrylamide.

**Fig. 5.** Representative MS/MS data of a chymotryptic peptide. The amino acid sequence from 255 to 266 of SU(A) contains two Asn residues. MS data of this peptide yielded only a 1 Da increment, indicating that only one of the two Asn was converted to Asp upon the PNGase F treatment. The MS/MS data show that Asn196 and not Asn193 was converted to Asp. “y” and “b” ions correspond to peptide fragmentation products generated from the C and N terminus of the peptide, respectively. The molecular weight difference between y7 and y8 corresponded to acrylamide modified cysteine (depicted as C* in the figure).
various envelope subgroups interact specifically with one particular receptor, Tva? It is possible that Tva was selected, evolutionarily, as a receptor for ASLV, because it interacts with conserved elements of ASLV envelope. It is likely that an ancestral ASLV envelope, which presumably recognized one receptor, gave rise to the current diverse set of envelopes that recognize several distinct receptors. This means that the parental envelope must have acquired the ability to recognize additional receptors. Such an event would be favored if there was some interaction between the basic envelope structure and the novel receptor. Mutations in the variable regions of the envelope that play a specific role in receptor recognition (hr1, hr2, vr3) would enhance any basic interaction and bring about the high-affinity interactions that are characteristic of the interactions of a retroviral envelope and its receptor. This type of scenario suggests, when envelope evolves the ability to use a new receptor, that not all the proteins on the surface of the host are equally likely to be chosen as targets. Rather, there would be a strong probability that the protein chosen as a receptor would be a protein for which the parental envelope would already have a modest affinity, which could explain why a carbohydrate that appears to be conserved would interact with a specific receptor.

There are alternative possibilities: Delos et al. (2002) suggested that the presence of carbohydrate plays a critical role in properly positioning one of the variable loops of SU(A). When Delos et al. (2002) mutated the S or T residues just downstream of the glycosylated asparagines glycosylation sites one at a time, three sites had a major effect on protein expression and folding (N59, N196, and N254). Of these, only the N254 mutant produced enough protein to assay; the mutant protein bound Tva poorly, and Delos et al. (2002) suggested that the carbohydrate at this position might be important, although this site is relatively far [on the linear amino acid sequence of SU(A)] from any of the variable regions associated with binding to specific receptors (hr1, hr2, vr3). If proximity to a variable region is important in the proper positioning of one of the variable segments that interact with specific receptors, N196 is quite close to hr2. It is also possible that one of the functions of the carbohydrate is to shield the charged amino acids (originally Asn, converted to Asp by PNGase F) to which the carbohydrate is linked. In such a model, the exposure of charged residues on the surface of SU(A) could affect its interaction with Tva. This type of model would suggest that the function(s) of the carbohydrate on SU(A) do not depend on any single site having carbohydrate modification but rather on the modification of several of the carbohydrate groups on SU(A). If this is the case, the loss of any one individual carbohydrate group would not be critical; however, removing several of the carbohydrate groups could significantly impair Tva binding.

Materials and methods

Vector constructions and cell culture

SU(A)-rlgG is a fusion of the Schmidt–Ruppin A Rous sarcoma virus SU glycoprotein with a portion of the constant region of a rabbit IgG (Zingler and Young, 1996). The protein was produced from TF/SU(A)-19 cells, a clonal DF-1 chicken cell line that constitutively expresses SU(A)-rlgG, as described previously (Holmen and Feder-
A soluble form of the chicken ASLV(A) Tva receptor, sTva-mlG, a fusion of the extracellular domain of the chicken Tva receptor with a portion of the constant region of a mouse IgG (Holmen et al., 2001), was produced from TF/cksTva-15 cells, a clonal DF-1 cell line that constitutively expresses chicken sTva-mlG, as described previously (Melder et al., 2003). Both the TF/SU(A)-19 and TF/cksTva-15 cells were routinely grown in Dulbecco’s modified Eagle’s medium [DMEM (GIBCO/BRL)] supplemented with 10% fetal bovine serum (GIBCO/BRL), 100 units of penicillin per milliliter and 100 μg of streptomycin per milliliter (Quality Biological, Inc., Gaithersburg, MD), and 250 μg of G418 per milliliter (GIBCO/BRL) at 39 °C and 5% CO₂.

Immunoprecipitations and pull-down assays

The SU(A)-rlG protein was immunoprecipitated from a 500-μl aliquot of TF/SU(A)-19 cell culture supernatant with 50 ml of anti-rabbit IgG-agarose beads (Sigma) for ≥1 h at 4 °C. The chicken sTva-mlG protein was immunoprecipitated from a 500-μl aliquot of TF/cksTva-15 cell culture supernatant with 50 μl of anti-mouse IgG-agarose beads (Sigma) for ≥1 h at 4 °C. The protein–antibody agarose bead complexes were collected by centrifugation and washed twice in dilution buffer [50 mM Tris–Cl, pH 8.3, 150 mM NaCl] and reconcentrated three times. The concentrated and washed sample was then applied to a 1.0-ml Protein A column (Pharmacia) equilibrated with 50 mM Tris, pH 8.3, 150 mM NaCl) and reconcentrated two times. The concentrated sample was diluted 5-fold with dilution buffer (50 mM Tris, pH 8.3, 150 mM NaCl) and reconcentrated two times. The concentrated and washed sample was then applied to a 1.0-ml Protein A column (Pharmacia) equilibrated with dilution buffer. The loaded column was washed with 30 ml dilution buffer followed by 30 ml of 10 mM Tris, pH 8.0. The sample was eluted from the column with 100 mM glycine, pH 3.0, into tubes that contained 100 μl 2 M Tris, pH 8.3. Peak fractions were pooled and concentrated using two 3-ml 10K cutoff centrifugal concentration units (Filtron). Samples were diluted to 10 ml with 20 mM Hepes, pH 7.4, containing 500 mM NaCl, and reconcentrated three times. The sequence of the N terminus of the purified SU(A)-rlG was determined by the Mayo Clinic Protein Core Facility using Edman degradation performed with a Applied Biosystems cLC492 Protein Sequencer. Analysis of the purified SU(A)-rlG allowed the assignments of the first 11 amino acids: DVHLLEQPGNL, which corresponds to the expected sequence of mature ASLV(A) SU.

Western transfer analysis

The denatured immunoprecipitates were fractionated by SDS-PAGE (12% polyacrylamide) and transferred to a nitrocellulose membrane. The filters were blocked with 10% nonfat dry milk (NFDM) in PBS, and probed with 0.05 μg/ml peroxidase-conjugated goat anti-mouse IgG or 0.05 μg/ml peroxidase-conjugated goat anti-rabbit antibodies (Kirkegaard and Perry Laboratories, Gaithersburg, MD) in rinse buffer (10 mM Tris–Cl, pH 8.0, 100 mM NaCl, 1 mM EDTA, 0.1% Tween 20, and 1% NFDM), and washed in rinse buffer. Protein–antibody complexes were detected with the Western Blot Chemiluminescence Reagent (NEN) according to the manufacturer’s instructions. The immunoblot was then exposed to Kodak X-Omat film.

Expression and purification of SU(A)-rlG

TF/SU(A)-19 cells were grown to confluence in normal growth media. Twenty-four hours before the SU(A)-rlG protein was collected for purification, the media on confluent TF/SU(A)-19 cells was replaced with serum-free DMEM without supplements. The culture supernatants were harvested, filtered through a 0.45-μm filter, and stored at −80 °C. Fresh media was added to the cultures, and after an additional 24 h, supernatant was harvested a second time and processed as described above. Under these conditions, the cell supernatants contained an average of 650 ng SU(A)-rlG per milliliter.

Five liters of cleared supernatant was concentrated approximately 50-fold by filtration on a 30-kDa cutoff tangential flow membrane (Vivaflow). The concentrated sample was diluted 5-fold with dilution buffer (50 mM Tris, pH 8.3, 150 mM NaCl) and reconcentrated two times. The concentrated and washed sample was then applied to a 1.0-ml Protein A column (Pharmacia) equilibrated with dilution buffer. The loaded column was washed with 30 ml dilution buffer followed by 30 ml of 10 mM Tris, pH 8.0. The sample was eluted from the column with 100 mM glycine, pH 3.0, into tubes that contained 100 μl 2 M Tris, pH 8.3. Peak fractions were pooled and concentrated using two 3-ml 10K cutoff centrifugal concentration units (Filtron). Samples were diluted to 10 ml with 20 mM Hepes, pH 7.4, containing 500 mM NaCl, and reconcentrated three times. The sequence of the N terminus of the purified SU(A)-rlG was determined by the Mayo Clinic Protein Core Facility using Edman degradation performed with a Applied Biosystems cLC492 Protein Sequencer. Analysis of the purified SU(A)-rlG allowed the assignments of the first 11 amino acids: DVHLLEQPGNL, which corresponds to the expected sequence of mature ASLV(A) SU.

Reduction and acrylamide modification of cysteines

The cysteine modification was performed according to a procedure described previously (Brune, 1992). Briefly, SU(A)-rlG was incubated in 0.3 M Tris–HCl, pH 8.3, containing 2% SDS and 0.1 M DTT at 65 °C for 30 min. The mixture was cooled to 37 °C and incubated with 2 M acrylamide for 30 min. The modified protein was precipitated with cold acetone and dried under vacuum during centrifugation. At this stage, samples were divided into two parts. One part was directly subjected to HPLC purification...
while another part was deglycosylated before the HPLC purification step.

**Deglycosylation of SU(A)-rIgG with PNGase F**

Modified SU(A)-rIgG was resuspended in 50 mM Na₃PO₄, pH 7.5, containing 0.1% SDS and 2% β-mercaptoethanol and denatured at 95 °C for 5 min. The samples were allowed to cool to 37 °C, then Triton X-100 (1% final concentration) and 500 Units of PNGase F (New England Biolabs) were added to the sample and the reaction was incubated at 37 °C for 1 h. The deglycosylated protein was precipitated with cold acetone and dried under vacuum during centrifugation.

**HPLC purification of proteins**

Glycosylated and deglycosylated protein samples were purified separately following the procedure described below. The dried samples were first suspended in 80 μl of 6 M guanidinium hydrochloride solution at 37 °C for 30 min. Subsequently, 20 μl of an aqueous solution of 2% acetonitrile and 1% formic acid was added and the mixture was centrifuged at 5000 × g for 10 min. The supernatant was then injected into a phenyl reverse phase column (TOSOH). The protein was eluted with a 40-ml linear gradient of 2–90% buffer at 37. The proteolytic digestions contained cold acetone and dried under vacuum during centrifugation.

**Proteolytic digestion of SU(A)-rIgG**

HPLC purified SU(A)-rIgG containing fractions were pooled and desiccated. The proteolytic digestions contained 5 μg of SU(A)-rIgG and 0.5 μg of one of the following three proteases: trypsin, chymotrypsin, or GluC (Roche). Trypsin digestion was carried out in 50 mM ammonium-bicarbonate buffer at 37 °C for 16 h. Chymotryptic digestion of the protein was performed in 10 mM Tris–HCl buffer, pH 7.8, at 25 °C for 16 h. GluC digestion was carried out in 50 mM ammonium-acetate, pH 4.0, at 25 °C for 16 h. The proteolytic fragments were dried under vacuum during centrifugation and resuspended in aqueous solution of 1% trifluoroacetic acid.

**Mass spectrometric analyses**

MS analyses yield accurate molecular weight measurements of proteins and peptides, while MS/MS analyses provide amino acid sequence information for small-molecular-weight peptides (MW less than 4 kDa) based on internal fragmentation of peptide bonds. MS spectra were recorded using MALDI-ToF or quadrupole time-of-flight (Q-ToF) techniques. MALDI-ToF experiments were performed with Kratos Axima-CFR instrument (Kratos Analytical Instruments, Manchester, UK). HPLC purified fractions of glycosylated and deglycosylated SU(A)-rIgG were mixed with equal volumes of a saturated sinapinic acid solution in 50% acetonitrile in water and applied on the MALDI sample plate. The proteolytic peptide fragments of SU(A)-rIgG were analyzed using α-cyano-4-hydroxy-cinematic as a matrix.

MS and MS/MS analyses of the proteolytic peptides were carried out using a Micromass (Manchester, UK) Q-ToF-II instrument equipped with an electrospray source and a Micromass cap-LC. Peptides were separated with the Thermo Hypersil Keystone 72105-030315 precolumn (Thermo Environmental, Franklin, MA) and the Micro-Tech Scientific (Vista, CA) ZC-10-C18SBWX-150 column using two sequential linear gradients of 5–40% acetonitrile for 35 min and 40–90% acetonitrile for 10 min. MS/MS analysis data and the MASCOT search engine, www.matrixscience.com were used to identify SU(A) peaks from the NCBIINr primary sequence database.

**Acknowledgments**

We thank Hilda Marusiodis for help in preparing the manuscript. This work was supported in part by grants to M.J.F. from NIH AI48682 and the Mayo Foundation, and work in S.H.H.’s laboratory was supported by the National Cancer Institute and NIGMS.

**References**


